



苏州大学

SOOCHOW UNIVERSITY



# 基于图像识别算法的物体认知 系统关键技术研究及应用

答辩人：周靖越

导师：姚望舒

计算机科学与技术 嵌入式与物联网

计算机科学与技术学院



1

研究背景及技术基础

2

可增量图像识别框架

3

多层神经网络推理过程切分算法

4

物体认知系统设计及实现

5

低资源嵌入式设备上的实现

6

总结和展望

7

研究成果



# 第1章 研究背景及技术基础

**1.1** 研究背景

**1.2** 存在问题

**1.3** 技术基础



**物体认知**是指以嵌入式计算机为核心，利用传感器采集各类标识信息对物体进行识别的技术。

射频识别



超声波



声音







苏州大学

SOOCHOW UNIVERSITY

1.1

# 研究背景



**图像数据**具有设备适用性强，应用范围广等优势

设备性能

图像采集

数据传输

处理算法



身份识别  
车牌识别  
缺陷检测





## 1.2

### 存在问题

一、系统对持续学习的能力的需求。

二、终端设备无法提供充足资源。

三、设备间数据传输具有限制。

## 1.3

### 技术基础

物联网框架

云计算  
边缘计算

深度学习

迁移学习  
增量学习



# 第2章 可增量图像识别框架

**2.1** 效果展示

**2.2** 框架及具体实现

**2.3** 测试实验





苏州大学

SOOCHOW UNIVERSITY

2.1

效果展示



帮助

98ms

diaper: 0.18  
sleeping bag: 0.17  
plastic bag: 0.07

每帧图像处理时间  
帮助界面调出按钮

帮助

124ms

plastic bag: 0.25  
quill: 0.07  
sleeping bag: 0.07

一、以增量的形式进行训练。

二、少数据量，快速的训练。

当前模式，以及当前模式下已处理的图像数

模型推理得到的结果

推理模式与训练模式切换按钮

用户提供的标签

推理模式，第2998张  
无

模式切换

训练模式，第315  
餐巾纸

餐巾纸

帮助 91ms  
mouse: 0.52  
candle: 0.03  
plane: 0.03

推理模式，第117张  
无

模式切换



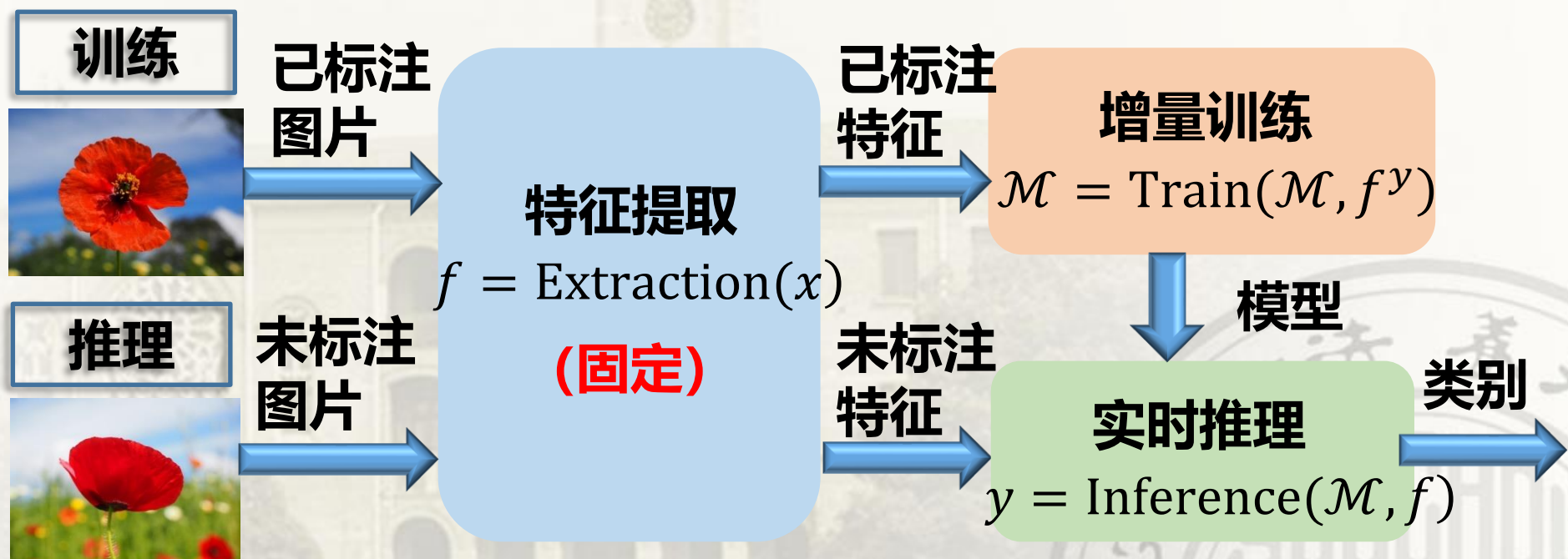


一、以增量的形式进行训练。

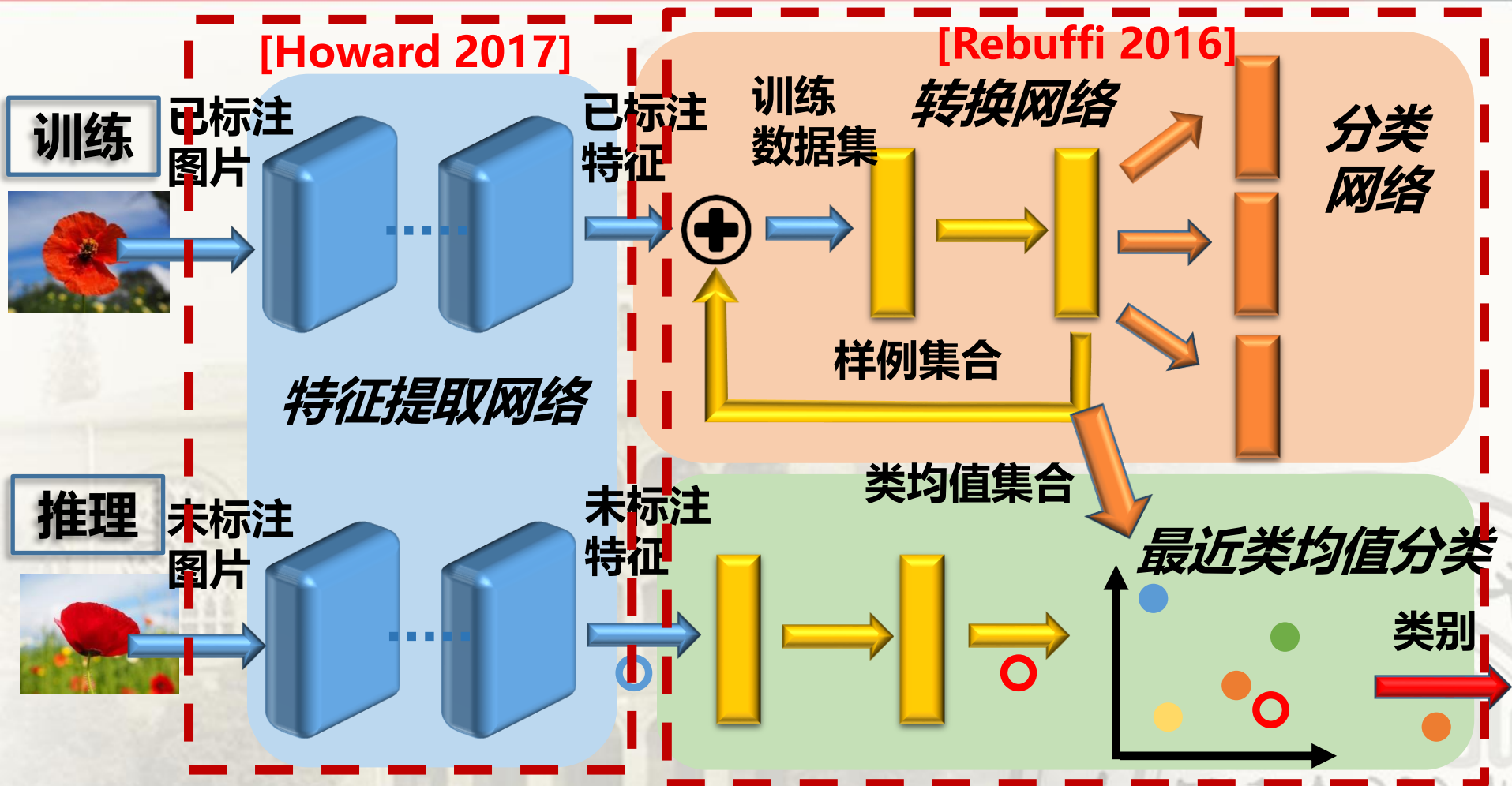
使用增量的训练算法

二、少数据量，快速的训练。

固定的特征提取模块

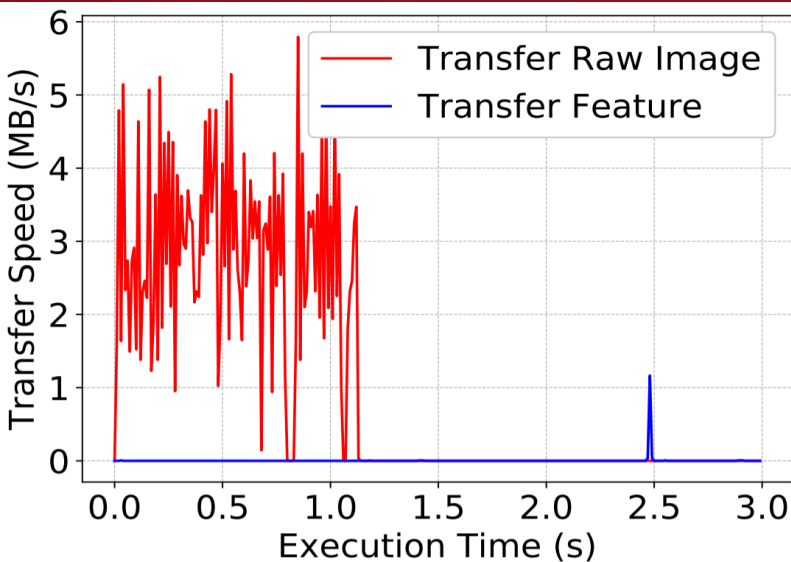


与物联网环境相关的优势：特征提取模块位于终端设备，部署之后不必更新，可压缩数据，减少传输量，增加安全性与隐私性。



[Howard 2017] Howard A G, Zhu M, Chen B, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications [EB/OL]. <https://arxiv.org/abs/1704.04861>, 17 Apr 2017.

[Rebuffi 2016] Rebuffi S A, Kolesnikov A, Sperl G, et al. iCaRL: Incremental Classifier and Representation Learning [A]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition [C]. Piscataway, NJ: IEEE, 2017: 2001-2010.



(a) 特征提取**压缩率与速度**。

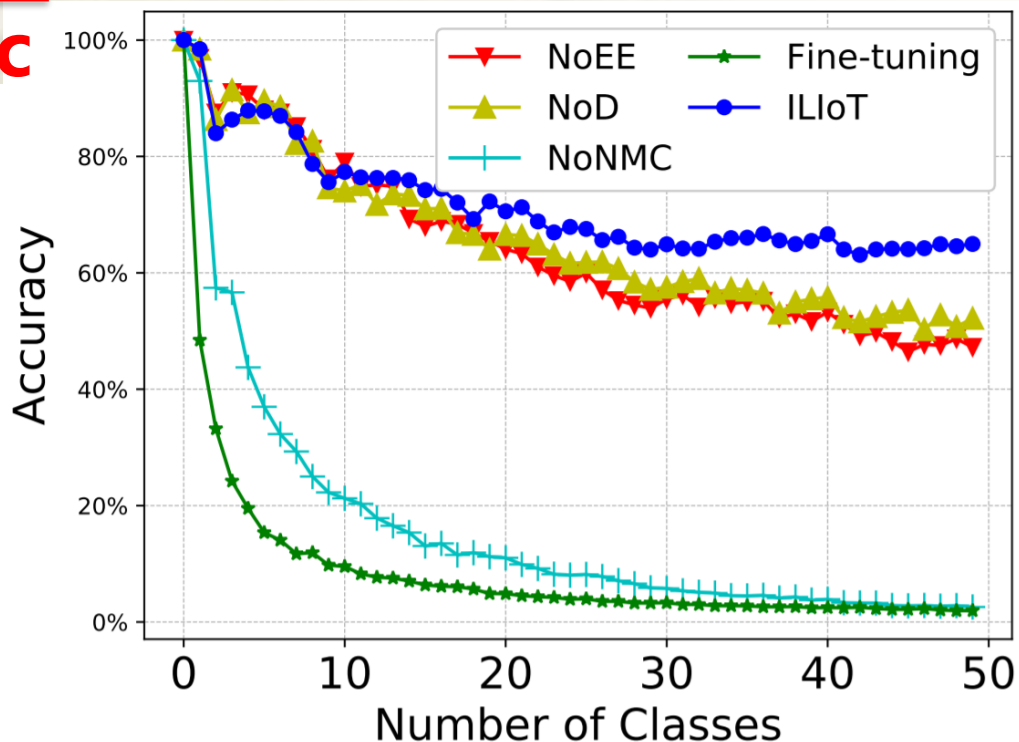
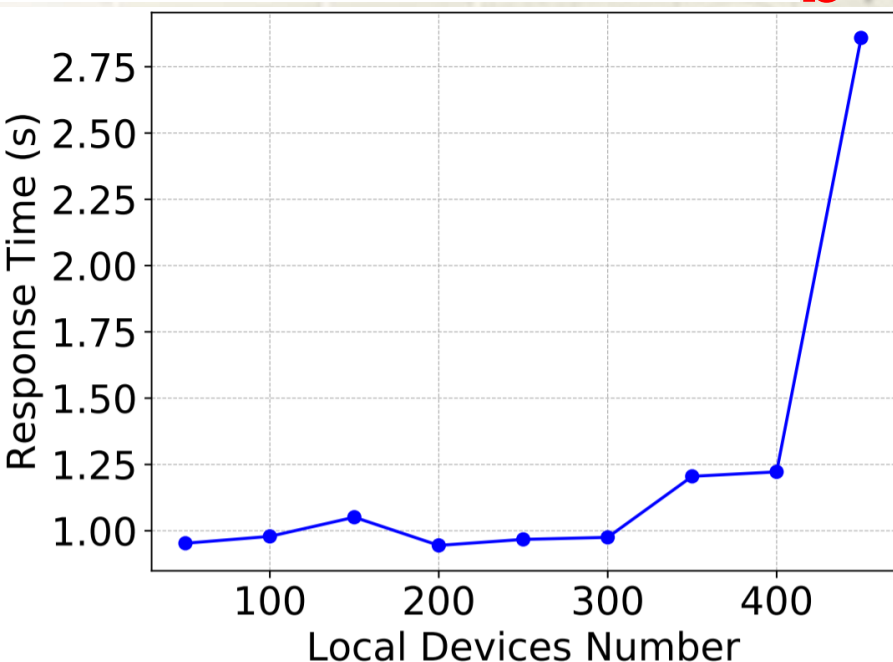
设备：树莓派，图像大小：2.98MB

(b) 推理运算的**并发容量**测试。

设备：树莓派，请求间隔：6分钟

(c) 分类**准确度**测试。

数据集：Core50





# 第3章 多层神经网络推理过程 切分算法

**3.1** 建立模型

**3.2** 切分算法

**3.3** 测试实验

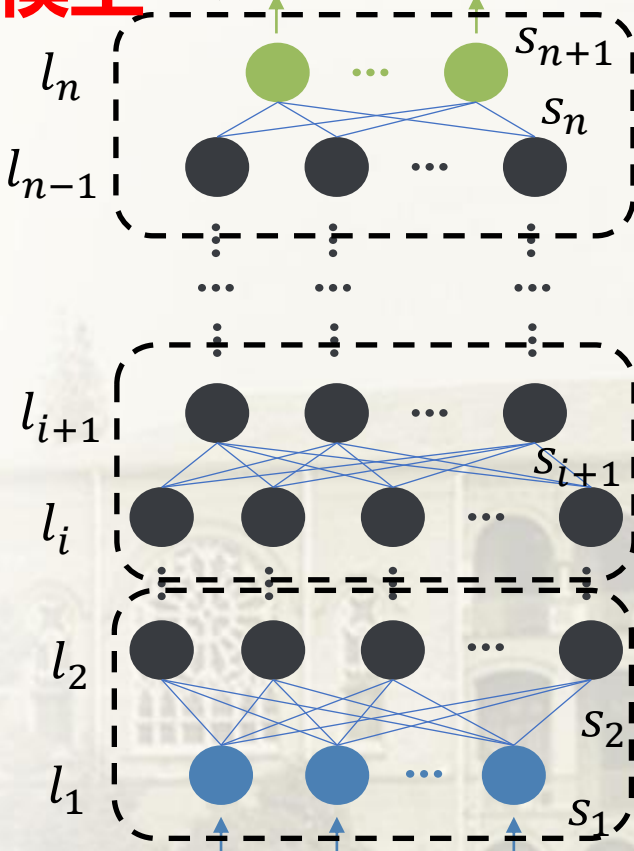
该研究内容已录用于 Zhou J, Wang Y, Ota K, et al. AAIoT: Accelerating Artificial Intelligence in IoT Systems [J]. IEEE Wireless Communications Letters, 2019.



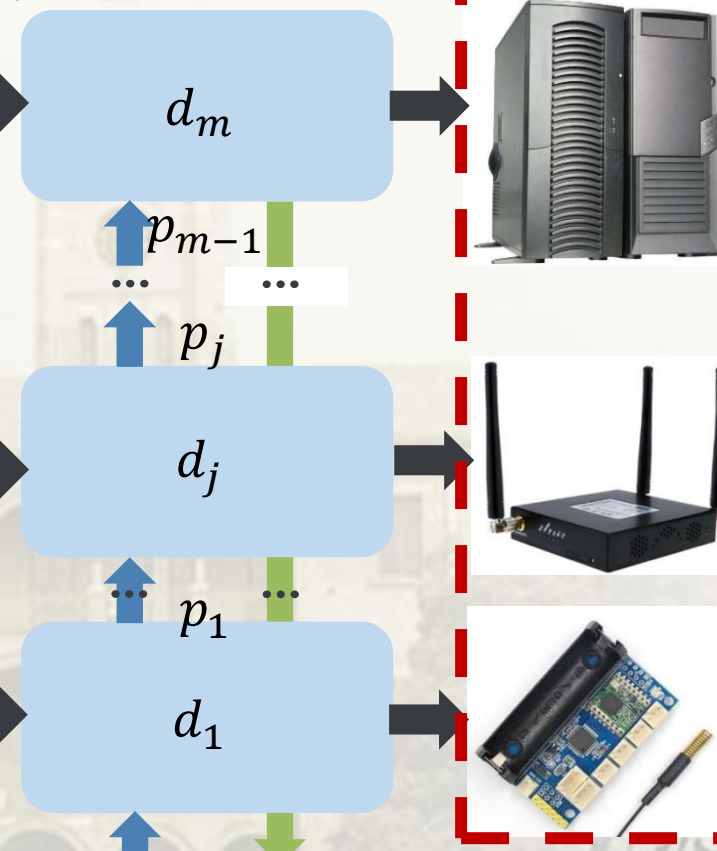


# 模型

## 多层神经网络



## 多层物联网结构



神经网络层:  $l$   
 数据传输量:  $s$   
 传输速度:  $p$   
 物联网设备:  $d$   
 运算时间:  $c$   
 分配策略:  $a$

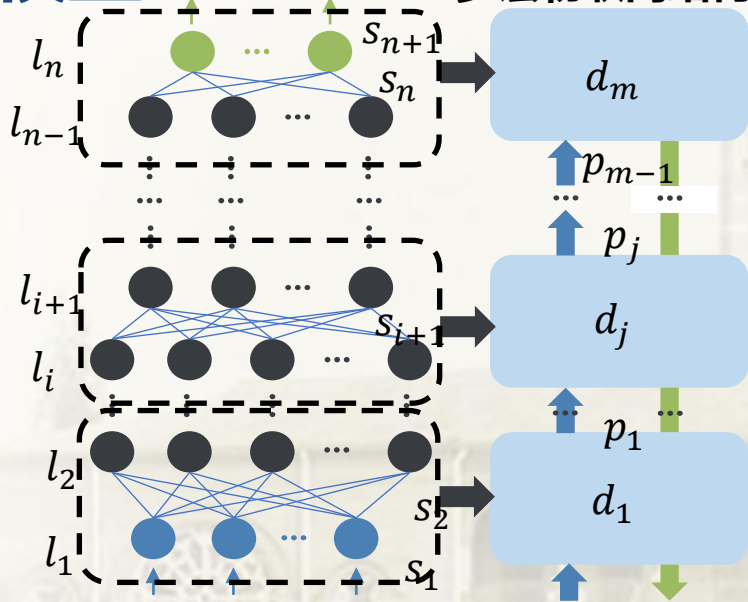
# 评价标准

响应时间 = 运算时间 + 传输时间

$$= \text{calcu} + \text{trans} \left\{ \begin{aligned} \text{calcu} &= \sum_{i=1}^n c_i^{a_i} \\ \text{trans} &= \sum_{i=1}^n \sum_{j=a_{(i-1)}}^{(a_i)-1} (s_i + s_{n+1}) / p_j \quad (a_0 = 1) \end{aligned} \right.$$



## 模型 多层神经网络 多层物联网结构



## 1. 状态变量设计与计算

输入：数据传输量： $s$  输出：状态变量： $h$   
 传输速度： $p$   
 运算时间： $c$

$$h_i^j = \begin{cases} 0 & i = 1, j = 1 \\ h_1^{j-1} + (s_1 + s_{n+1}) / p_{j-1} & i = 1, j > 1 \\ h_{i-1}^1 + c_{i-1}^1 & i > 1, j = 1 \\ \min(h_{i-1}^j + c_{i-1}^j, h_i^{j-1} + (s_i + s_{n+1}) / p_{j-1}) & i > 1, j > 1 \end{cases}$$

$h_i^j$  代表了满足将网络层  $l_i$  的输入数据传输到设备  $d_j$  的条件时所能达到的最短响应时间

## 2.

## 获取分配策略

算法：取得分配策略

Input：状态变量  $H = \{h_1^1, \dots, h_i^j, \dots, h_{n+1}^m\}$

Output：分配策略  $A = \{a_1, a_2, \dots, a_n\}$

$a_n = \arg \min_{j \in 1, \dots, m} h_{n+1}^j, t = a_n$

for  $i \in n-1, \dots, 1$  do

while  $h_{i+1}^t \neq h_i^t + c_i^t$  do  $t = t - 1$

$a_i = t$

return  $A$

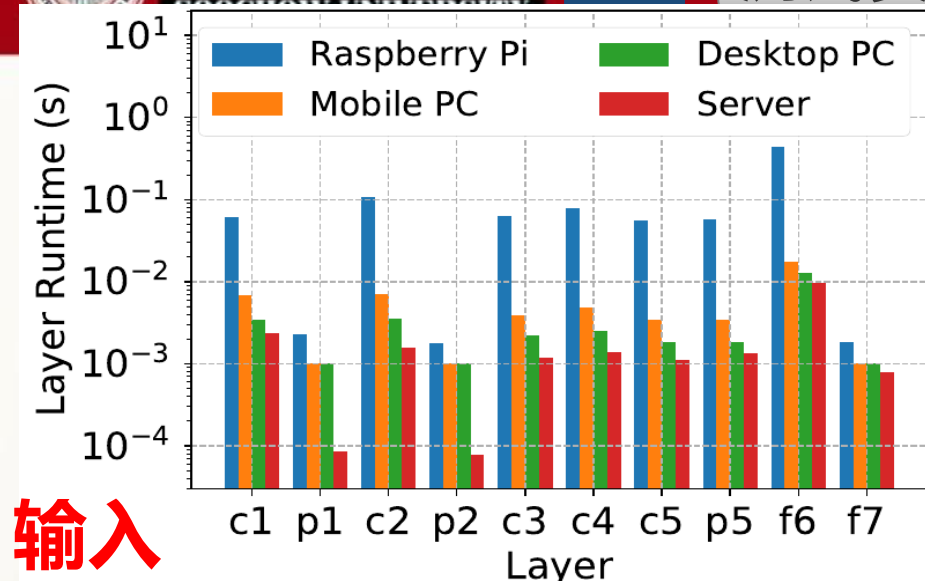
## 算法分析

时间复杂度： $(mn)$

$n$  为神经网络层数

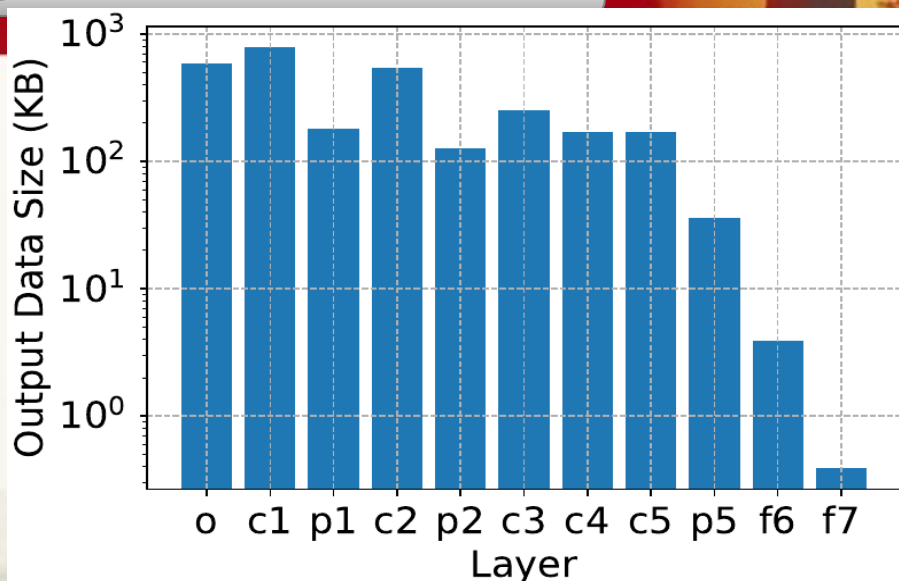
$m$  为设备层数

能取得最优解

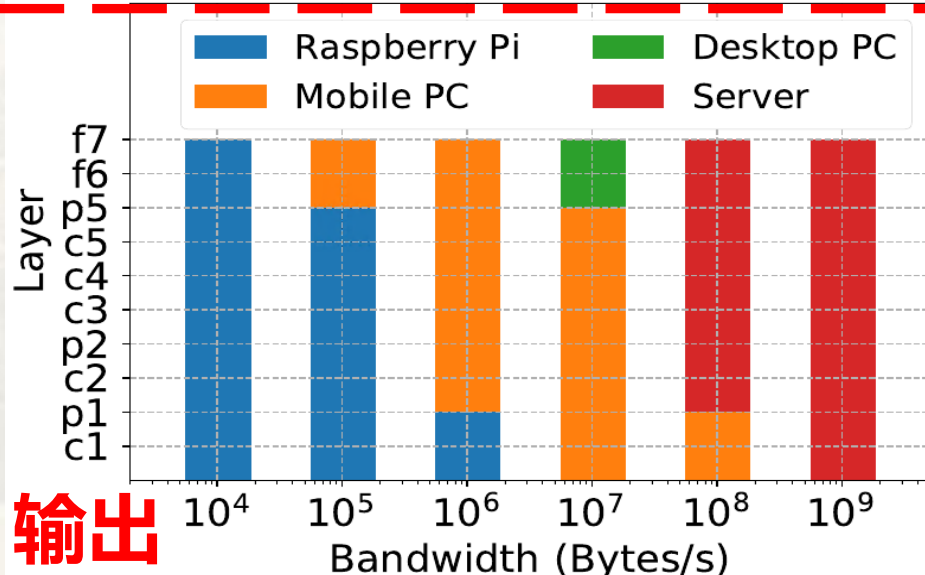


输入

(a) 测量得到的运算时间

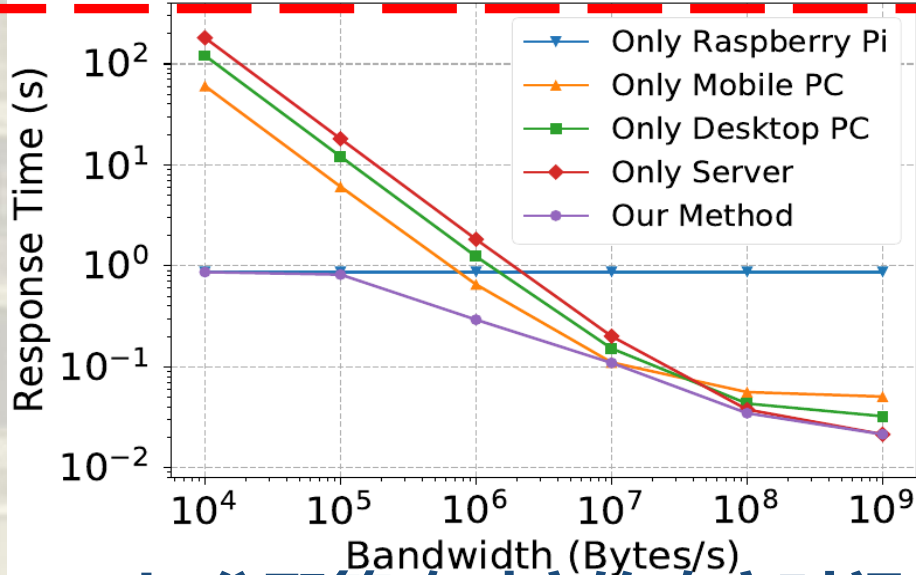


(b) 测量得到的数据传输量



输出

(c) 算法输出的分配策略



(d) 分配策略对应的响应时间



# 第4章 物体认知系统设计与实现

**4.1** 系统设计与实现

**4.2** 适用性展示

**4.3** 系统资源占用分析

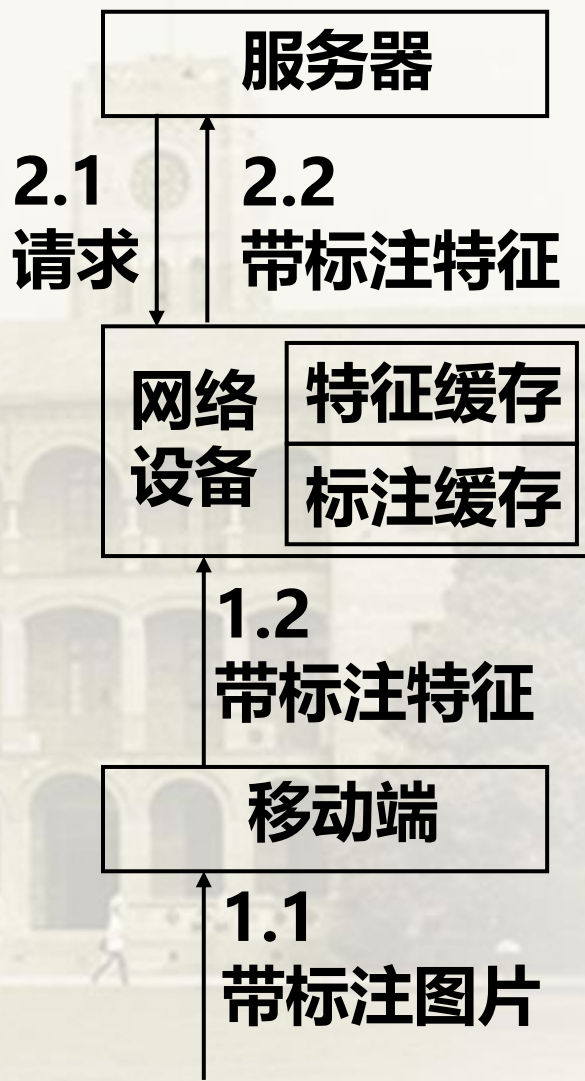




## 推理过程



## 训练过程



## 台式机

部署了**增量训练与实时推理**模块。

## 华为云服务器

具有固定的网络地址，是网络访问的服务端。

## 手机

部署了**特征提取**模块，并负责与外界交互。



苏州大学

SOOCHOW UNIVERSITY

4.2

适用性展示

18

基础物体  
识别能力  
连续学习八  
个类别的物  
体后的分类  
效果。

帮助

121ms  
hair slide: 0.22  
screwdriver: 0.06  
lighter: 0.05



推理模式, 第39张  
U盘

模式切换

帮助

140ms  
ladle: 0.28  
hammer: 0.15  
rifle: 0.10



推理模式, 第97张  
笔

模式切换

帮助

85ms  
modem: 0.13  
cassette: 0.08  
shoe shop: 0.08

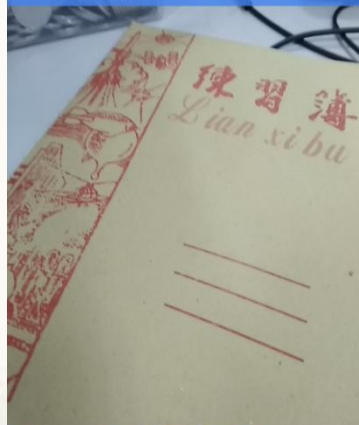


推理模式, 第160张  
手机

模式切换

帮助

131ms  
envelope: 0.83  
handkerchief: 0.01  
binder: 0.01



推理模式, 第223张  
练习本

模式切换

帮助

107ms  
cup: 0.49  
coffee mug: 0.31  
mixing bowl: 0.05



推理模式, 第307张  
杯子

模式切换

帮助

101ms  
cocktail shaker: 0.22  
espresso maker: 0.14  
oil filter: 0.08



推理模式, 第382张  
瓶子

模式切换

帮助

121ms  
mouse: 0.85  
frying pan: 0.10  
modem: 0.02



推理模式, 第486张  
鼠标

模式切换

帮助

105ms  
ladle: 0.78  
bobsled: 0.02  
spatula: 0.02



推理模式, 第627张  
勺子

模式切换



苏州大学

SOOCHOW UNIVERSITY


4.2

适用性展示

## 适应多种任务场景

先部署后训练的模式，以及极低的训练成本使用户得以对每一个特定的任务**建立自己独有的模式**。

帮助 142ms  
photocopier: 0.14  
printer: 0.08  
ballpoint: 0.07



推理模式, 第49张  
玻璃

模式切换

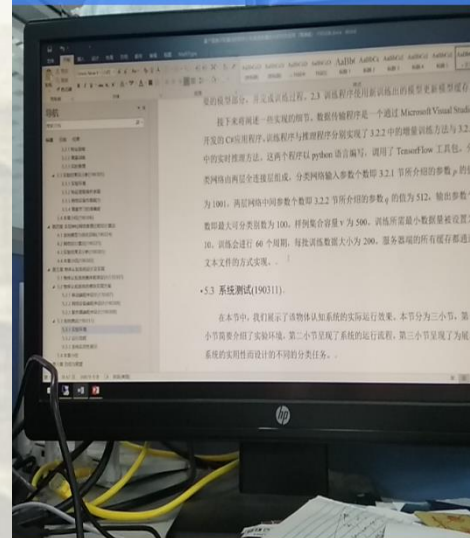
帮助 103ms  
lab coat: 0.44  
paper towel: 0.15  
plastic bag: 0.15



推理模式, 第115张  
纸张

模式切换

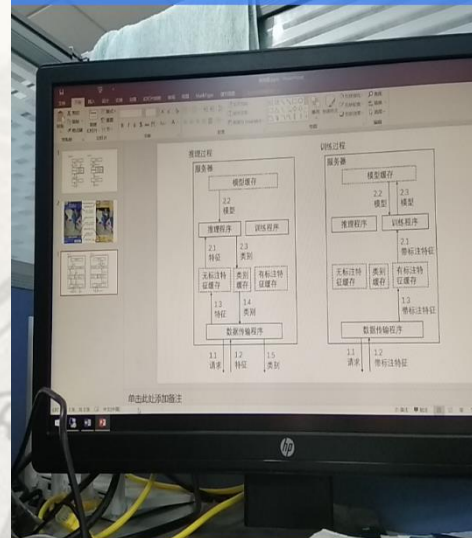
帮助 81ms  
monitor: 0.30  
slide rule: 0.19  
web site: 0.16



推理模式, 第174张  
word

模式切换

帮助 100ms  
monitor: 0.22  
photocopier: 0.16  
desktop computer: 0.08



推理模式, 第247张  
ppt

模式切换

(a) 垃圾分类

(b) 应用识别





## 适应多种任务场景

训练的权力下放给用户，用户可以**自由控制**训练的程度，以及细化分类的方向。

帮助 92ms  
traffic light: 0.17  
switch: 0.15  
nail: 0.05

1

推理模式, 第2394张  
1

1 模式切换

帮助 109ms  
parachute: 0.20  
tick: 0.10  
hook: 0.07

2

推理模式, 第2352张  
2

1 模式切换

(c) 数字识别

帮助 105ms  
puck: 0.44  
bottlecap: 0.22  
thimble: 0.10



推理模式, 第323张  
1

1 模式切换

帮助 105ms  
bottlecap: 0.56  
puck: 0.15  
lens cap: 0.08



推理模式, 第266张  
2

1 模式切换

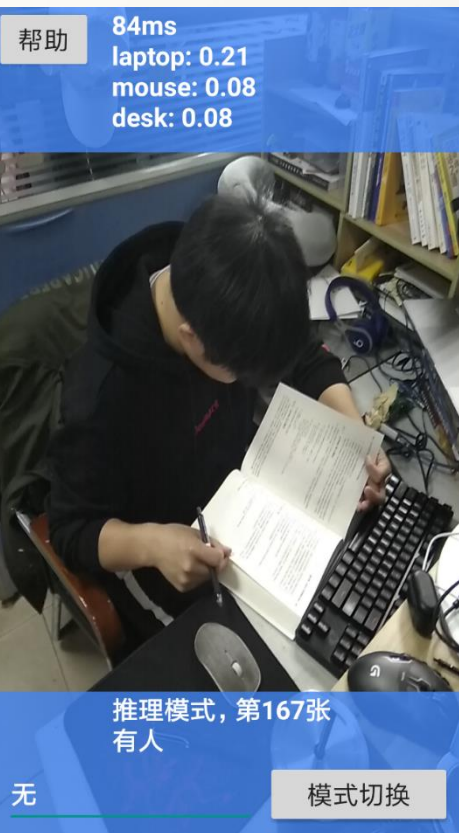
(d) 数量识别



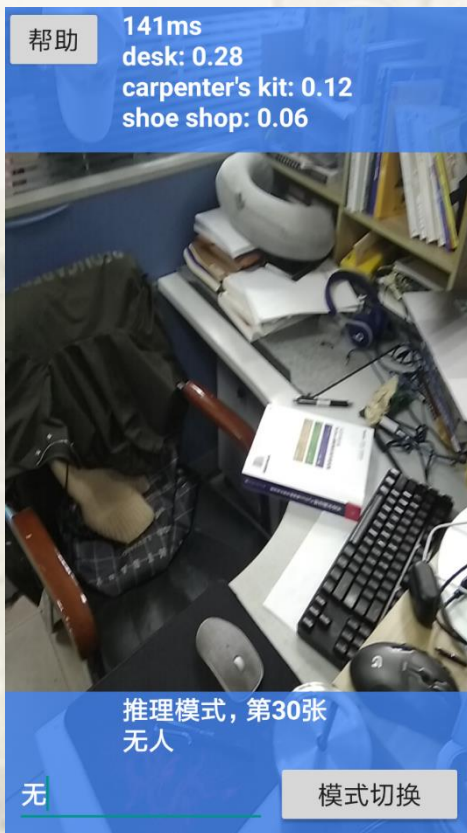


## 适应多种任务场景

由用户主导应用的功能设计，具有**广泛的适用性**，做到以同一个程序适应多种任务场景。



(e) 是否有人



(f) 是否开灯





## 结论

手机，树莓派等终端设备足以支持基于深度学习的图像识别推理过程。

### (a) 常见图像识别模型的资源占用情况

模型	存储占用	推理运算次数
VGG	527 MB	--
AlexNet	240 MB	--
MobileNet	0.47 M-4.24 M (个参数)	14 M-569 M (次乘加运算)
SqueezeNet	0.5 MB	--

### (b) 常见设备所拥有的资源

设备	内核架构	处理速度	内存	存储空间
红米Note 4X	Cortex-A53	最高 2.0 GHz	3 GB	32 GB
树莓派3B	Cortex-A53	1.2 GHz	1 GB	16 GB
S32K144	Cortex-M4	最高 112 MHz	60 KB	512 KB
MKL25Z128	Cortex-M0+	最高 48 MHz	16 KB	128 KB



## 结论

常见的嵌入式微控制器**无法支持**基于深度学习的图像识别推理过程。

### (a) 常见图像识别模型的资源占用情况

模型	存储占用	推理运算次数
VGG	527 MB	--
AlexNet	240 MB	--
MobileNet	0.47 M-4.24 M (个参数)	14 M-569 M (次乘加运算)
SqueezeNet	<b>0.5 MB</b>	--

### (b) 常见设备所拥有的资源

设备	内核架构	处理速度	内存	存储空间
红米Note 4X	Cortex-A53	最高 2.0 GHz	3 GB	32 GB
树莓派3B	Cortex-A53	1.2 GHz	1 GB	16 GB
<b>S32K144</b>	Cortex-M4	最高 112 MHz	<b>60 KB</b>	512 KB
MKL25Z128	Cortex-M0+	最高 48 MHz	16 KB	128 KB



# 第5章 低资源嵌入式设备上的实现

- 5.1 算法设计以及测试
- 5.2 软硬件设计以及实现
- 5.3 系统测试以及性能分析





训练

已标注  
图片



已标注  
特征

增量训练

$$\mathcal{M} = \text{Train}(\mathcal{M}, f^y)$$

框架

推理

未标注  
图片



未标注  
特征

模型

实时推理

$$y = \text{Inference}(\mathcal{M}, f)$$

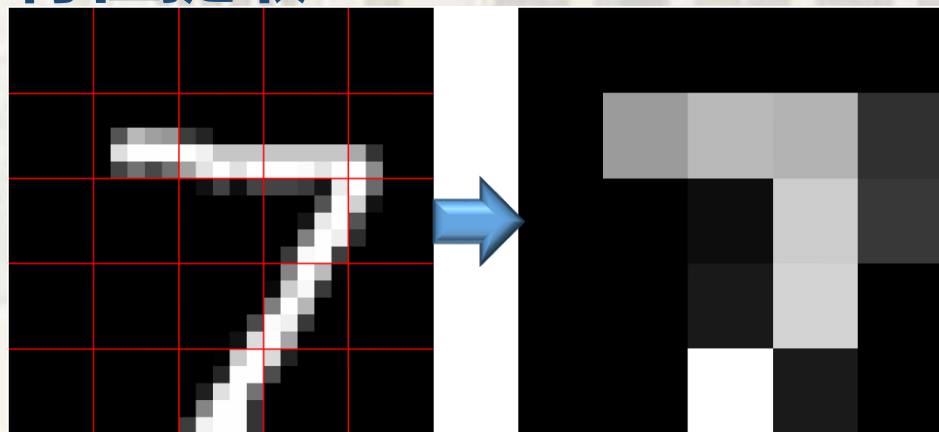
类别

特征提取

$$f = \text{Extraction}(x)$$

(固定)

特征提取



增量训练

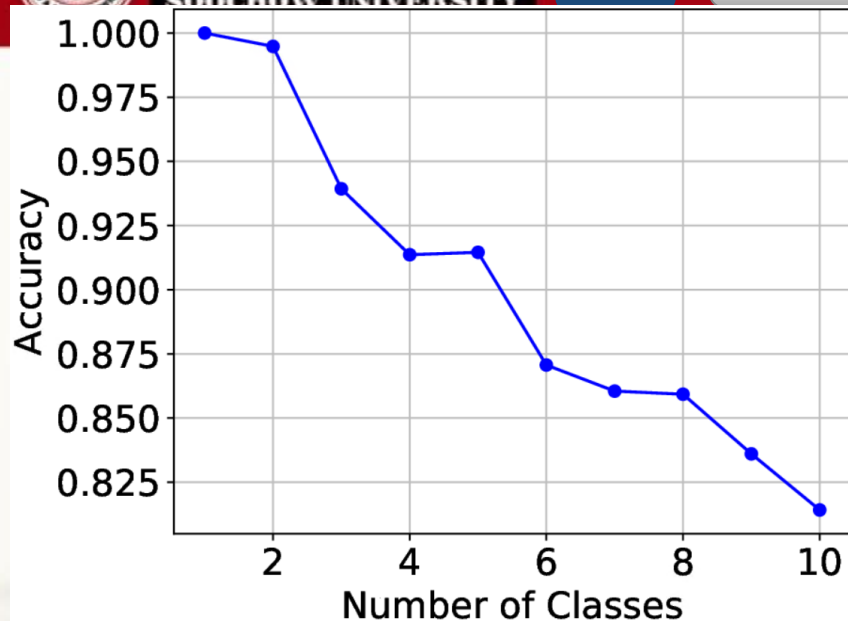
$$m^y = m^y (r-1) / r + f^y / r$$

实现

实时推理

$$y = \arg \min_{i \in 1, \dots, n} (\|m^i - f\|)$$

学习速率:  $r$  模型:  $= \{m^1, \dots, m^n\}$



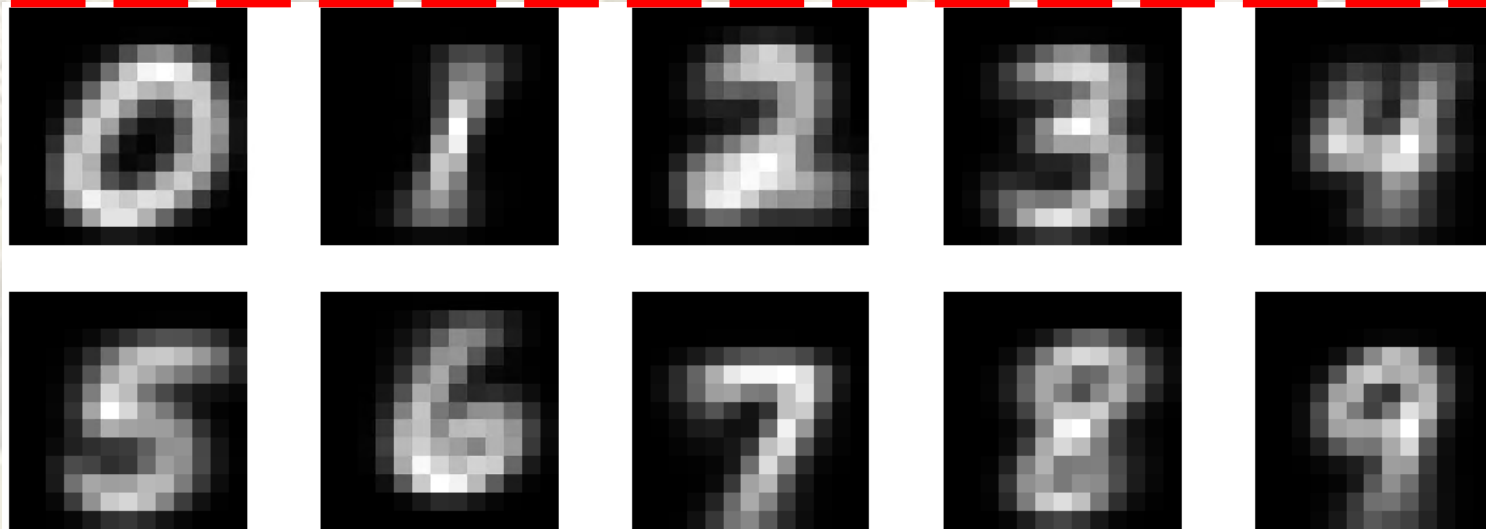
(a) MNIST上的准确率

在MNIST手写数字集上  
达到**0.814**的分类准确率

实验参数

特征  $f$  的大小为  $13*13$

学习速率  $r$  为 1000



(b) 训练得到的模型



苏州大学

SOOCHOW UNIVERSITY

5.2

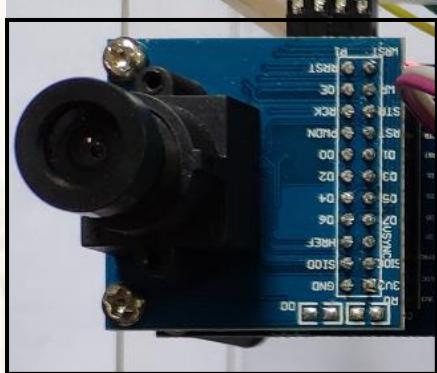
软硬件设计及实现



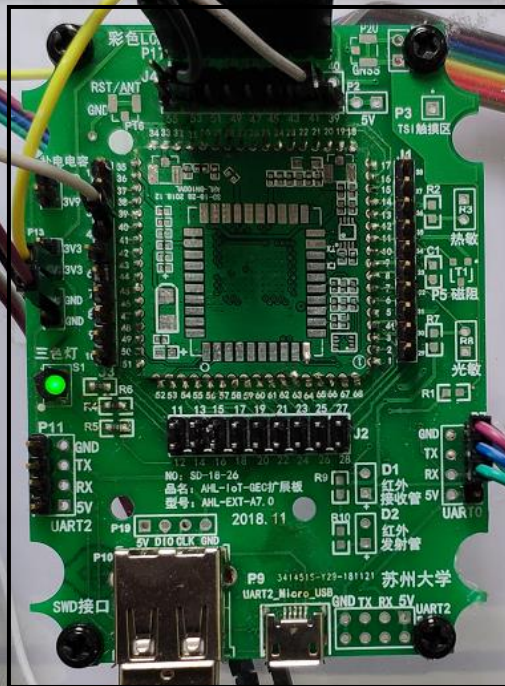
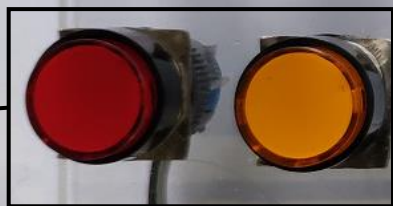
# 硬件设计

选用**S32K144**作为运算核心

摄像头

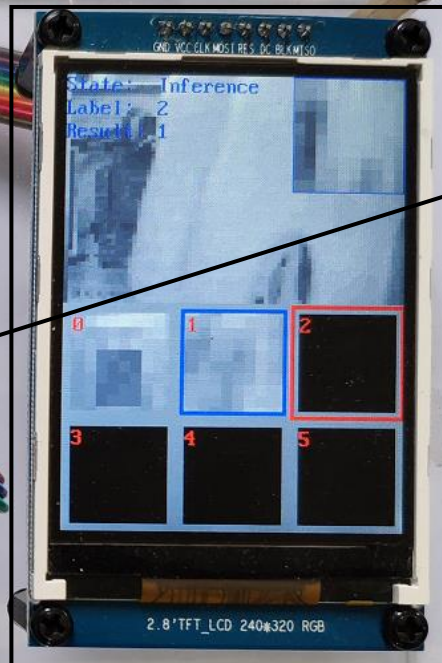


按钮



主控板

液晶屏



基于图像识别算法的物体认知系统

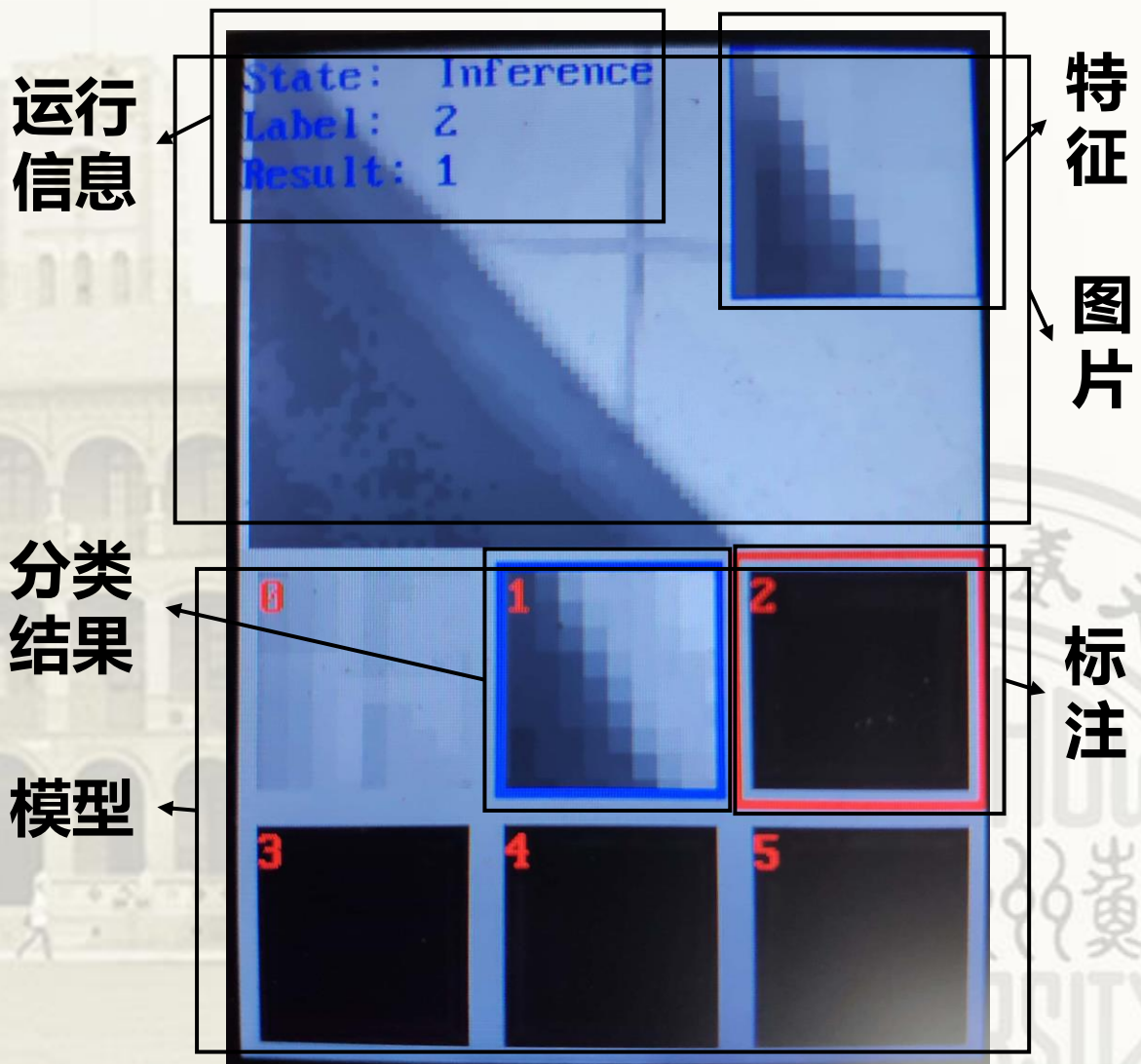




# 运行流程



# 显示界面







## 系统性能

内存占用: **7.6 KB**

存储占用: **13.0 KB**

推理帧率: **0.61 Hz**

训练帧率: **0.59 Hz**

### (a) 常见图像识别模型的资源占用情况

模型	存储占用	推理运算次数
VGG	527 MB	--
AlexNet	240 MB	--
MobileNet	0.47 M-4.24 M (个参数)	14 M-569 M (次乘加运算)
SqueezeNet	<b>0.5 MB</b>	--

### (b) 常见设备所拥有的资源

设备	内核架构	处理速度	内存	存储空间
红米Note 4X	Cortex-A53	最高 2.0 GHz	3 GB	32 GB
树莓派3B	Cortex-A53	1.2 GHz	1 GB	16 GB
<b>S32K144</b>	Cortex-M4	最高 112 MHz	<b>60 KB</b>	512 KB
MKL25Z128	Cortex-M0+	最高 48 MHz	16 KB	128 KB



# 第6章 总结与展望

**6.1** 总结

**6.2** 展望



## 6.1

### 总结

一、可增量图像识别框架。

二、多层神经网络推理过程切分算法。

三、移动端与嵌入式设备上的实现。

## 6.2

### 展望

框架

提高准确率，物体定位等其他能力

切分算法

考虑功耗，输入数据测量方法

实现

多用户处理，性能优化



苏州大学

SOOCHOW UNIVERSITY



# 第7章 研究成果







## 已录用论文

[1] **Zhou J**, Wang Y, Ota K, et al. AAIoT: Accelerating Artificial Intelligence in IoT Systems [J]. IEEE Wireless Communications Letters, 2019. **(SCI二区)**

## 其他

- [1] 王宜怀, 张建, 刘辉. 窄带物联网NB-IoT应用开发共性技术 [M]. 北京: 电子工业出版社, 2019. (参与撰写)
- [2] 参加横向科研项目P11801617 《燃气表软硬件设计》.





苏州大学

SOOCHOW UNIVERSITY



**敬请各位老师批评指正！**

